

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-259507

(43) 公開日 平成11年(1999) 9月24日

(51) Int.Cl.⁸

識別記号

F I

G 0 6 F 17/30

G 0 6 F 15/40

3 7 0 D

G 0 6 T 1/00

H 0 4 N 5/76

B

H 0 4 N 5/76

G 0 6 F 15/403

3 5 0 C

5/91

15/66

4 7 0

H 0 4 N 5/91

N

審査請求 未請求 請求項の数3 O L (全 6 頁)

(21) 出願番号

特願平10-60835

(22) 出願日

平成10年(1998) 3月12日

(71) 出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72) 発明者 谷口 行信

東京都新宿区西新宿三丁目19番2号 日本

電信電話株式会社内

(72) 発明者 外村 佳伸

東京都新宿区西新宿三丁目19番2号 日本

電信電話株式会社内

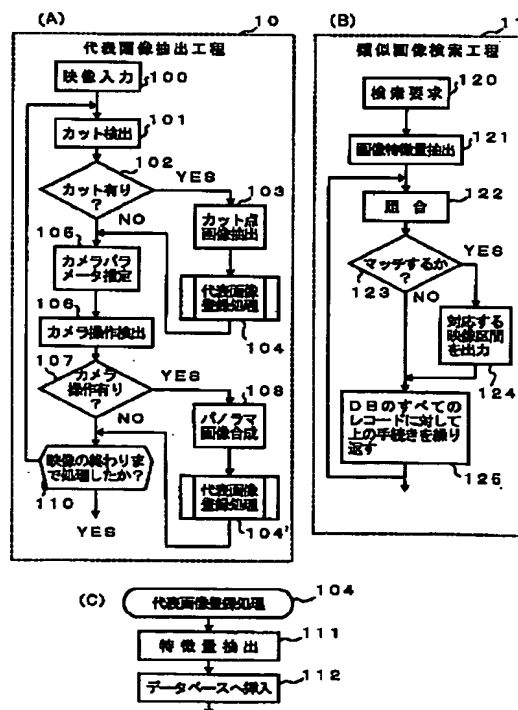
(74) 代理人 弁理士 小笠原 吉義 (外1名)

(54) 【発明の名称】 映像検索方法および映像検索プログラムを格納した記録媒体

(57) 【要約】

【課題】 映像から抽出した代表画像によって映像を検索するシステムにおいて、カメラ操作を含む映像に対してもシーン全体の画像内容を表し、かつ、空間的位置関係を保存した代表画像を抽出できるようにする。また、時間的な検索を可能とすることを目的とする。

【解決手段】 検索対象の映像から参照画像と類似した代表画像を検索する類似画像検索工程11で用いるための代表画像を抽出する代表画像抽出工程10では、カメラ操作を含む画像列を抽出し、抽出した画像列を合成することによってパノラマ画像を合成する処理を行うことにより、シーン全体の画像内容を表した代表画像を自動的に抽出する。また、カメラの動きに関する情報を代表画像と関連付けて保存しておくことにより、時間的な検索条件を空間的な検索条件に置き換えて、パノラマ画像として合成された代表画像に関する時間的な検索を可能とする。



1

【特許請求の範囲】

【請求項1】 映像から複数枚の代表画像を抽出する代表画像抽出工程と、参照画像と類似した代表画像を検索する類似画像検索工程とを具備する映像検索方法であって、前記代表画像抽出工程が、カメラ操作を含む画像列を抽出する工程と、抽出した画像列を合成することによってパノラマ画像を合成する工程を含むことを特徴とする映像検索方法。

【請求項2】 請求項1記載の映像検索方法において、前記代表画像抽出工程においてカメラの動きに関する情報を代表画像と関連付けて保存しておき、前記類似画像検索工程において時間的検索条件を空間的検索条件に置き換えることによって検索を行うことを特徴とする映像検索方法。

【請求項3】 映像から複数枚の代表画像を抽出する代表画像抽出工程と、参照画像と類似した代表画像を検索する類似画像検索工程とを具備する映像検索方法をコンピュータによって実現するためのプログラムを格納した記録媒体であって、前記代表画像抽出工程においてカメラ操作を含む画像列を抽出する工程と、抽出した画像列を合成することによってパノラマ画像を合成する工程とを含む処理を、コンピュータに実行させるプログラムを格納したことを特徴とする映像検索プログラムを格納した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、映像データベースシステム等における映像検索方法および映像検索プログラムを格納した記録媒体に関するものである。

【0002】

【従来の技術】キーワードなどの文字情報ではなく、画像特徴（色ヒストグラム、エッジ特徴など）に基づいた画像検索システムが開発されている（M.Flicker, H.Sawhney, W.Niblack, J.Ashley, Q.Huang, B.Dom, M.Gorkani, J.Hafner, D.Lee, D.Petkovic, D.Steele, P.Yanker, Query by Image and Video Content: The QBIC System, IEEE Computer Magazine, Vol.28, No.9, pp.23-32, 1997）。

【0003】文字情報に基づいたキーワード検索は、画像データベースにおいても重要な検索方法であるが、これはキーワード付与に人手を要すること、画像は人によってあるいは状況によって様々な意味を持つ（多義性）ためキーワード選択が困難であることが問題であった。これらの問題点を解決するために、画像特徴に基づいた検索方式が望まれていた。映像データベースにおける検索方法にも同様の問題点があり、画像特徴に基づいた検索方式が望まれていた。

【0004】利用者がどのような方法で検索要求をシステムに与えるかには、次のようなバリエーションがあり得る。

2

(a) 第1の方法は、利用者が一枚あるいは複数枚の画像を与えて、その画像と類似した画像を提示するように要求するものである。

【0005】(b) 第2の方法は、検索要求文を利用者が与えて、その文を何らかの知識を用いてシステムが扱いやすい要求条件に変換することによって検索を行うものである。例えば「山と青空の写っているシーンが欲しい」という利用者の検索要求に対して、システムは「山」は緑色で「青空」は青色であり、「山」の領域は「空」の領域の下にあるという知識を用いて、「緑色の領域の上に青色の領域が存在する画像を検索せよ」というシステムにとって扱い易いものに変換し、予め蓄積されているインデクスと検索要求を照合し合致する画像を提示する。

【0006】映像は画像が時間的に並んだものであるので、以上説明したような画像検索技術を応用することで映像検索も可能であると考えられていた。しかし、映像には膨大な数の画像が含まれているので、実用的な検索を行うためには前段階として、映像の中から重要な画像（代表画像と呼ぶ）を、少数抽出しておき、代表画像についてだけ画像特徴量を算出し、インデクスとするという方式がある（M.Flicker, H.Sawhney, W.Niblack, J.Ashley, Q.Huang, B.Dom, M.Gorkani, J.Hafner, D.Lee, D.Petkovic, D.Steele, P.Yanker, Query by Image and Video Content: The QBIC System, IEEE Computer Magazine, Vol.28, No.9, pp.23-32, 1997）。

【0007】代表画像の抽出方法としては、映像の場面のvarietyであるカット点を検出し、カット点の画像を代表画像とする方法があった。

【0008】

【発明が解決しようとする課題】しかし、上記従来技術には、カメラを動かしながら撮影された映像については適切な代表画像を抜き出すことが困難であるという問題点があった。

【0009】図6を用いてこの問題点について説明する。図6に示すシーン（池、山、空の映っているシーン）を、カメラを下から上に動かしながら撮影した映像（画像61, 62, 63, 64, 65）を考える。カット点を検出する従来の方法で抽出される代表画像は、シーンの先頭画像61である。この代表画像には、山が映っていないので「山の写ったシーンを提示せよ」という検索要求に応えることができない。

【0010】この問題を解決するために、今度は画像61, 62, 63, 64, 65の5枚の画像を代表画像として抽出した場合を考える。これらの代表画像を画像データベースに蓄えておけば、「山の写っているシーンを提示せよ」という検索要求に応えることができる。しかし、代表画像を抽出した時点で代表画像の間の位置関係が失われているので、「池の上に山の写っているシーンを提示せよ」といった検索要求に応えることはできな

い。

【0011】結局、以上のような従来の方法では、シーン全体の画像内容を表し、かつ、その空間的位置関係を保存した代表画像を抽出できないという問題点があった。また、映像を扱う場合には、時間的な検索を行えることが望まれる。例えば、利用者が2枚の画像A、Bを与えて、「Aの後にBが現れるシーンを探せ」といった検索要求に応えるようにしたいわけである。従来技術では、代表画像を抽出した時点で時間的な情報は失われていたので、時間的な検索を行えなかった。

【0012】本発明は、上記問題点を解決するためになされたものであり、カメラ操作を含む映像に対してもシーン全体の画像内容を表し、かつ、空間的位置関係を保存した代表画像を抽出でき、また、時間的な検索を可能とすることを目的とする。

【0013】

【課題を解決するための手段】本発明は、映像から複数枚の代表画像を抽出する代表画像抽出工程と、参照画像と類似した代表画像を検索する類似画像検索工程とを具備する映像検索方法であって、前記代表画像抽出工程が、カメラ操作を含む画像列を抽出する工程と、抽出した画像列を合成することによってパノラマ画像を合成する工程を含むことを特徴とする。カメラ操作を含む画像列を抽出し、その画像列からパノラマ画像を合成するようにしたこと、シーン全体の画像内容を表した代表画像を自動的に抽出できるようになる。

【0014】さらに第2の発明は、前記代表画像抽出工程においてカメラの動きに関する情報を代表画像と関連付けて保存しておき、前記類似画像検索工程において時間的な検索条件を空間的な検索条件に置き換えることによって検索を行うことを特徴とする。カメラの動きに関する情報を代表画像と関連付けて保存しておくことにより、時間的な検索条件を空間的な検索条件に置き換えることができるようになる。

【0015】以上の各処理工程をコンピュータによって実現するためのプログラムは、コンピュータが読み取り可能な可搬媒体メモリ、半導体メモリ、ハードディスクなどの適当な記録媒体に格納することができる。

【0016】

【発明の実施の形態】本発明の実施の形態について図を参照して説明する。

<実施の形態1>図1は本発明の実施の形態の処理フロー図である。大きく二つの工程、(a) 代表画像抽出工程、(b) 類似画像検索工程、に分けられる。

【0017】(a) 代表画像抽出工程図1(A)の点線で囲まれている部分が代表画像抽出工程10である。この工程10は、新たな映像をデータベース（以下、DBと呼ぶ）に登録する際に実行される。ステップ100では、映像データを入力する。具体的には、映像信号をデジタル化したり、他の媒体に記録されているデジタル映

像データを読み込む処理であったりする。ステップ101～110は、入力映像データを先頭から解析し、カット点とカメラ操作を検出することによって複数の代表画像を抽出する処理である。以下、具体的な手続きについて説明する。

【0018】ステップ101では、カット点を検出するための処理を行う。カット点検出の方法は、様々なものが開示されているが、単純に、現在着目している画像とその一つ前の画像の間の輝度差分をとり、その差分の絶対値和があるしきい値以上である場合にカット点であると判定することができる。ステップ102で、カットありと判定された場合には、ステップ103において現在入力されているカット点画像を代表画像として、ステップ104で代表画像登録処理手続きに渡す。

【0019】ステップ104の代表画像登録手続きについて説明する。図1(C)のステップ111では、登録しようとしている代表画像から、画像特徴量を抽出する。画像特徴量としては、様々なものを利用することができるが、例えば画像の輝度の平均値、色ヒストグラムなどを用いることができる。

【0020】ここでは、色ヒストグラムを例にとりて説明する。代表画像の各画素点 i のRGB値を R_i , G_i , B_i ($0 \leq R_i, G_i, B_i \leq 255$)とする。RGB値(R_i, G_i, B_i)は、もともと8ビットあるが、それぞれを2ビットに圧縮／合成して6ビット値 V_i に変換する。ヒストグラム $H(v)$, $v=0, 1, \dots, 63$ は、 $V_i=v$ となる画素数をカウントしたものであり、 $H(v)$ は正の整数値をとる。ステップ112では、代表画像データ、特徴量、映像の先頭からの時間などをデータベース(DB)の一つのレコードとして登録する。なお、画像特徴量としてヒストグラム $H(v)$ を登録する場合には、画像サイズに依存しないよう正規化してからレコードに登録する。

【0021】図2を用いて、具体的なDBの構成の例について説明する。DBの一つのレコードは、図2に示すように、レコードのID番号のフィールド21、画像データへのポインタのフィールド22、映像IDのフィールド23、映像区間のフィールド24、パノラマ画像か否かのフラグ25、カメラ動き情報のフィールド26から構成される。フィールド22には、この例では、画像データへのポインタとして画像ファイルのファイル名を格納しているが、もちろん画像データ自身を格納してもよい。フィールド23には、映像IDとしてオリジナルの映像に対するポインタ情報が格納される。フィールド24には、代表画像が対応する映像区間が格納される（0～10000は0秒から10000ミリ秒までの区間）。フィールド25は、代表画像のタイプを表すフラグが格納される。フィールド26については後述する。

【0022】ステップ105では、カメラパラメータを推定する。カメラパラメータの推定方法としては、様々

5

なものが開示されており、いずれを使ってもよいが、単純な例としては、平行移動モデルをカメラモデルとし、その最適パラメータをcoarse-to-fine戦略で探索するという方法がある(Taniguchi, Y. et al. Panorama Excerpts: Extracting and Packing Panoramas for Video Browsing, Proceedings of ACM Multimedia 97, pp.427-436.)。

【0023】ステップ106では、ステップ105で推定されたカメラパラメータに基づいてカメラ操作の有無を判定する。平行移動モデルの場合には、パラメータ値がすべて0であれば、カメラ操作がないと判定することができる。ノイズを考慮し、より厳密な判定を行うために、カメラモデルの当てはまりのよさを定量化する方法もある(Taniguchi, Y. et al. Panorama Excerpts: Extracting and Packing Panoramas for Video Browsing, Proceedings of ACM Multimedia 97, pp.427-436.)。

【0024】ステップ107で、カメラ操作ありと判定された場合には、ステップ108でパノラマ画像を合成する。より具体的には、カメラの動きをキャンセルするように画像列を重ねあわせることによって、背景に継ぎ目が見えないように画像を合成することができ、最終的に広視野のパノラマ画像を生成できる。抽出されたパノラマ画像はステップ104' (図1(C)の処理)で代表画像として登録される。

【0025】ステップ103では、カット点直後の画像を代表画像とする例について説明したが、カットから一定時間後の画像を代表画像とするようにしてもよい。また、カット点画像を代表画像としてすぐに登録するのではなく、次のカット点が発見される前にパノラマ画像を抽出したら、カット点画像は代表画像として登録しないようにすることも冗長な処理を省く意味で好適である。

【0026】(b) 類似画像検索工程

図1(B)の点線で囲まれている部分が類似画像検索工程11である。この工程11は、利用者が検索要求をシステムに与えるたびに実行される。

【0027】ステップ120では、利用者からの検索要求を入力する。ここでは、類似画像検索の例を考える。つまり、利用者は1枚の画像をシステムに与えて、検索結果として類似画像を含む映像区間を出力として得るというものである。したがって、ステップ120では、画像データが読み込まれる。

【0028】ステップ121では、利用者から与えられた画像から、画像特徴量を抽出する。どのような種類の特徴量を算出するかは検索条件によるが、例えば、色に関して類似しているものが欲しければ、色ヒストグラムなど色の特徴を反映したものとする。

【0029】ステップ122では、DBの一つのレコードに記録されている特徴量と、ステップ121で抽出された特徴量を照合し、距離を計算し、ステップ123の判定でその距離が許容値よりも小さい場合には、類似画

6

像が見つかったとして、ステップ124で対応する映像区間を出力する。色ヒストグラムの例では、利用者から与えられた画像から得られたヒストグラムをH1

(v)、DB内の代表画像から得られたヒストグラムをH2(v)とすれば、二つの画像の間の距離Dは、次式によって計算できる。

$$【0030】D = \sum |H1(v) - H2(v)|$$

(ただし、 \sum は $v=0$ から63までの総和)。ステップ122からステップ124までの処理を、DBに格納されているすべてのレコードについて繰り返す(ステップ125)ことによって、類似映像区間を列挙することができる。

【0031】<実施の形態2>次に、請求項2の発明に対応する第2の実施の形態について説明する。図3に、その処理フロー図を示す。本処理フロー図のほとんどの部分は、図1の対応する各ステップと同じであるので、ここでは、新規に加わったステップ301とステップ302についてのみ説明する。

【0032】ステップ301は、カメラ操作情報(カメラの動き情報)を抽出するステップであり、代表画像がカット点によって抽出されたものであれば、カメラは固定されていたとし、パノラマ画像である場合には、前述した図1(A)の代表画像抽出工程10におけるステップ105で推定されたカメラパラメータに対して、次に説明するような変換を施して、ステップ112でDBへ挿入する。

【0033】図4において、41~44は入力画像列を示し、それぞれ時刻 $t=0, 1, 2, 3$ の画像に対応している。この画像列から生成されるパノラマ画像が、画像45である。入力画像 I_t とパノラマ画像45の対応点(I_t の左上角の点)を、 P_t ($t=0, 1, 2, 3$)とする。対応点 P_t は、ステップ105で求めたカメラパラメータを足し合わせていくことで算出できる(平行移動モデルの場合)。この点列 P_t を、図2に示すDBのフィールド26にカメラ動き情報として格納しておく。

【0034】図3に戻って、類似画像検索工程11について説明する。本実施の形態では、検索要求として、利用者から2枚の画像IMG1、IMG2を受け取り、

「IMG1の類似画像のあとにIMG2の類似画像が現れる映像区間を検索せよ」というものを考える。

【0035】図3の類似画像検索工程11におけるステップ120では、2枚の画像IMG1とIMG2を入力する。ステップ121では、画像特徴量として上記実施の形態1と同様にして、IMG1、IMG2からそれぞれ色ヒストグラムH1、H2を算出する。ステップ122では、代表画像とIMG1、IMG2を色ヒストグラムH1、H2を用いて照合し、IMG1、IMG2の類似領域が同時に見つかったとき、ステップ302に処理を移す。この時点で、現在着目している映像区間には、

IMG1, IMG2の類似画像領域が存在することが分かったことになる。

【0036】ステップ302では、「IMG1のあとにIMG2が現れる」という時間条件を満足するかどうかを判定する。図4の例で、検索要求がIMG1＝

“A”，IMG2＝“E”である場合，を例にとつて説明する。ステップ122で，IMG1（“A”）とIMG2（“E”）に類似した領域46，47が見つかる。DBのフィールド26に設定されているカメラ動き情報を参照することにより，領域46は時刻 $t=0$ に，領域47は時刻 $t=3$ に現れることが分かる。したがって，上記時間条件を満たしていることが分かり，ステップ124で対応する映像区間を出力する。結局，時間的条件を含んだ映像検索が可能となる。

【0037】＜システムの構成例＞図5は，本発明を適用する映像検索システムの簡単な構成例を示している。処理装置1は，CPUおよびメモリ等からなり，図1に示す代表画像抽出工程10を実現するためのソフトウェアプログラム等によって構成される代表画像抽出部2，代表画像登録処理部3，また図1および図3に示す類似画像検索工程11を実現するためのソフトウェアプログラム等によって構成される類似画像検索部4を備える。これらを実現するプログラムは，例えばコンパクトディスク，フロッピーディスクなどの記録媒体から読み出してインストールすることができる。

【0038】代表画像抽出部2は，カメラから入力した映像信号をデジタル化したり，他の媒体に記録されている映像データ5を入力し，前述した処理により代表画像を抽出して，抽出した代表画像を代表画像登録処理部3により映像データ5とともに映像データベース6に登録する。

【0039】通信回線等で接続された利用者端末7から検索したい画像を指定した映像データの検索要求があるとき，類似画像検索部4は，前述した処理により映像デー

タベース6に登録された代表画像を用いて要求された映像データを検索し，検索結果を利用者端末7へ送信する。

【0040】

【発明の効果】本発明によれば，カメラ操作区間を自動的に検出し，パノラマ画像を合成し，代表画像とするようにしたので，カメラ操作を含む映像に対してもシーン全体の画像内容を表し，かつ，空間的位置関係を保存した代表画像を抽出できるようになるという効果がある。また，カメラ操作に関する情報を代表画像と関連付けておくことで，時間的な検索が可能となるという効果がある。

【図面の簡単な説明】

【図1】本発明の実施の形態の処理フロー図である。

【図2】データベースの構成について説明するための図である。

【図3】本発明の第2の実施の形態の処理フロー図である。

【図4】カメラ動き情報について説明するための図である。

【図5】本発明を適用する映像検索システムの構成例を示す図である。

【図6】従来技術の問題点について説明するための模式図である。

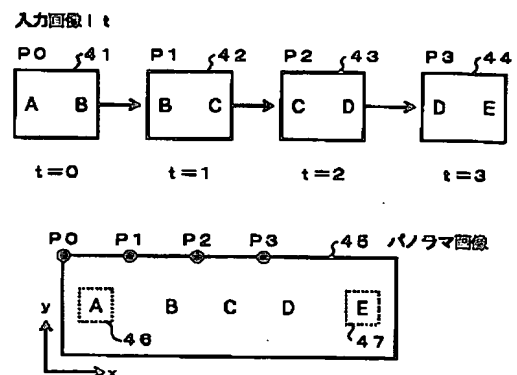
【符号の説明】

- 10 代表画像抽出工程
- 11 類似画像検索工程
- 100～110，301 代表画像抽出工程のステップ
- 120～125，302 類似画像検索工程のステップ
- 21～26 DBのフィールド
- 41～44 入力画像列
- 45 パノラマ画像
- 46，47 領域

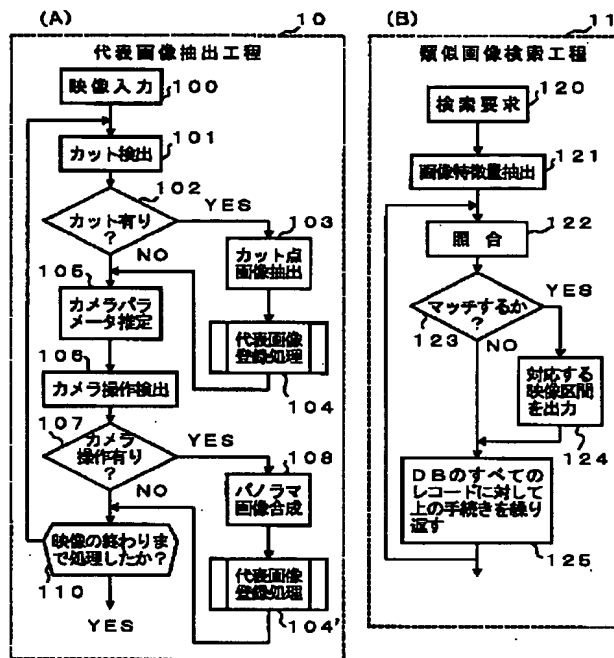
【図2】

{21}		{22}		{23}		{24}		{25}		{26}	
ID	画像データ	映像ID	区間	フラグ	FLAG	カメラ動き情報					
1	IMG001. jpg.	1	0～10,000	1		P0, P1, P2, ...					
2	IMG002. jpg.	1	10,000～15,000	0							
⋮	⋮	⋮	⋮	⋮	⋮	⋮					

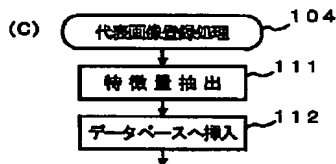
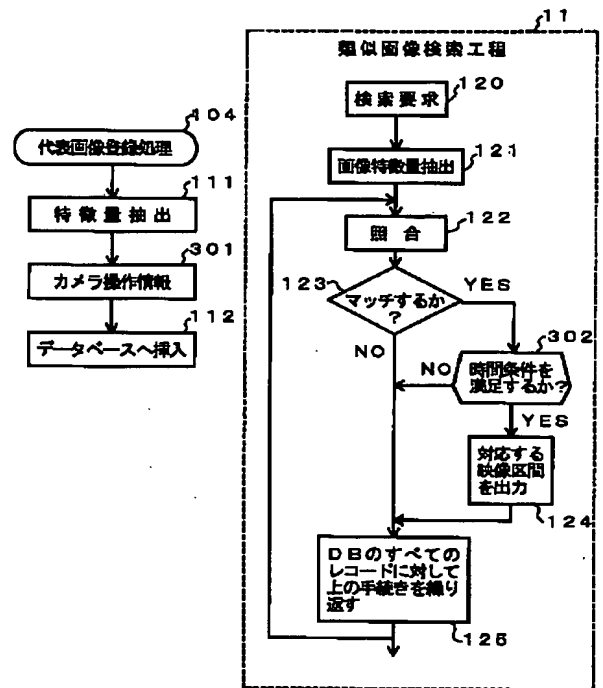
【図4】



【図1】

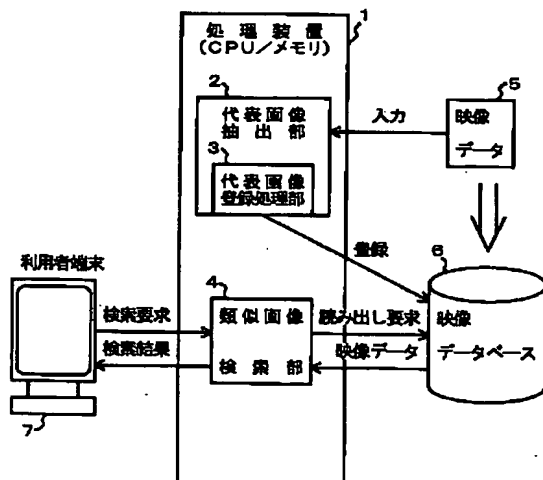


【図3】



【図5】

映像検索システムの構成例



【図6】

